

## Supplementary Info for:

# Common variants at five new loci associated with early-onset inflammatory bowel disease

Marcin Imielinski<sup>1,28</sup>, Robert N. Baldassano<sup>3,6,28</sup>, Anne Griffiths<sup>7,28</sup>, Richard K Russell<sup>8,28</sup>, Vito Annese<sup>14,28</sup>, Marla Dubinsky<sup>12,28</sup>, Subra Kugathasan<sup>22,28</sup>, Jonathan P. Bradfield<sup>1</sup>, Thomas D. Walters<sup>7</sup>, Patrick Sleiman<sup>1</sup>, Cecilia E. Kim<sup>1</sup>, Aleixo Muise<sup>7</sup>, Kai Wang<sup>1</sup>, Joseph P. Glessner<sup>1</sup>, Shehzad Saeed<sup>21</sup>, Haitao Zhang<sup>1</sup>, Edward C. Frackelton<sup>1</sup>, Cuiping Hou<sup>1</sup>, James H. Flory<sup>1</sup>, George Otieno<sup>1</sup>, Rosetta Chiavacci<sup>1</sup>, Robert Grundmeier<sup>4,6</sup>, Massimo Castro<sup>16</sup>, Anna Latiano<sup>16</sup>, Bruno Dallapiccola<sup>17</sup>, Joanne Stempak<sup>18</sup>, Debra J. Abrams<sup>3</sup>, Kent Taylor<sup>12</sup>, Dermot McGovern<sup>12</sup>, Western Regional Research Alliance for Pediatric IBD, International IBD Genetics Consortium, Melvin B. Heyman<sup>19</sup>, George D. Ferry<sup>20</sup>, Barbara Kirschner<sup>11</sup>, Jessica Lee<sup>25</sup>, Jonah Essers<sup>25</sup>, Richard Grand<sup>25</sup>, Michael Stephens<sup>23</sup>, Arie Levine<sup>14,24</sup>, David Piccoli<sup>3,6</sup>, Johan Van Limbergen<sup>9</sup>, Salvatore Cucchiara<sup>15</sup>, Dimitri S. Monos<sup>5</sup>, Stephen L. Guthery<sup>13</sup>, Lee Denson<sup>26</sup>, David C Wilson<sup>10</sup>, Struan F.A. Grant<sup>1,2,6</sup>, Mark Daly<sup>27</sup>, Mark S. Silverberg<sup>18,29</sup>, Jack Satsangi<sup>9,29</sup>, Hakon Hakonarson<sup>1,2,6,29</sup>

<sup>1</sup>Center for Applied Genomics, <sup>2</sup>Division of Human Genetics, <sup>3</sup>Division of Gastroenterology, <sup>4</sup>Center for Biomedical Informatics, and <sup>5</sup>Department of Pathology The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, 19104, USA, <sup>6</sup>Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, 19104, USA, <sup>7</sup>The Hospital for Sick Children, University of Toronto, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada, <sup>8</sup>Paediatric Gastroenterology and Nutrition, Royal Hospital for Sick Children, Glasgow, G3 8SJ, UK, <sup>9</sup>Gastrointestinal Unit, Division of Medical Sciences, School of Molecular and Clinical Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK, <sup>10</sup>Paediatric Gastroenterology and Nutrition, Royal Hospital for Sick Children, Edinburgh and Child Life and Health, University of Edinburgh, EH9 1UW, UK, <sup>11</sup>University of Chicago Comer Children's Hospital, Chicago, Illinois 60637, USA, <sup>12</sup>Departments of Pediatrics and Common Disease Genetics, Cedars Sinai Medical Center, Los Angeles CA, 90048, USA, <sup>13</sup>Department of Pediatrics, University of Utah School of Medicine and Primary Children's Medical Center, Salt Lake City, Utah, 84132, USA, <sup>14</sup>Units of Gastroenterology and Endoscopy, IRCCS-CSS Hospital, S. Giovanni Rotondo, Italy, <sup>15</sup>Pediatric Gastroenterology & Liver Unit, La Sapienza University of Rome & SIGENP, Italy, <sup>16</sup>Gastroenterology Unit, Bambino Gesù Children's Hospital, IRCCS, Rome & SIGENP, Italy, <sup>17</sup>Mendel Institute, La Sapienza University of Rome, IRCCS-CSS Hospital, S. Giovanni Rotondo, Italy, <sup>18</sup>Mount Sinai Hospital IBD Centre, University of Toronto, 441-600 University Avenue, Toronto, Ontario M5G 1X5, Canada, <sup>19</sup>Department of Pediatrics, University of California, San Francisco, San Francisco, 94143, CA, USA, <sup>20</sup>Department of Pediatrics, The Baylor College of Medicine, Houston, TX, 77030, USA, <sup>21</sup>Dept of Pediatrics, University of Birmingham, Alabama, 35294, USA, <sup>22</sup>Department of Pediatrics, Emory University School of Medicine and Children's Health Care of Atlanta, Atlanta, GA, 30322, USA, <sup>23</sup>Department of Pediatrics Medical College of Wisconsin, Milwaukee, WI, 53226, USA, <sup>24</sup>Pediatric Gastroenterology Unit, Wolfson Medical Center, Tel Aviv University, Israel, <sup>25</sup>Division of Gastroenterology, Children's Hospital, Boston, MA 02115 USA, <sup>26</sup>The Center for Inflammatory Bowel Disease, Division of Gastroenterology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, 45229, USA, <sup>27</sup>Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA, <sup>28</sup>These authors contributed equally to the manuscript. <sup>29</sup>These authors jointly directed the work.

## Supplementary Note

### Clinical data

All patients in the discovery (DC) and CHOP replication (RC1) cohorts were diagnosed prior to their 19<sup>th</sup> birthday and fulfilled standard IBD diagnostic criteria. Cohort characteristics are shown in **Supplementary Table 1**. Family history of IBD was obtained with focus on first degree relatives. A patient was considered to be of Jewish heritage when at least 2 grandparents were known to be Jewish. Phenotypic characterization was based on a modification of the Montreal classification such that the definitions of L1 & L3 were both extended to include disease within the small bowel proximal to the terminal ileum and distal to the ligament of Treitz. Disease above the ligament of Treitz was recorded separately; perianal disease included only those patients with perianal abscess and/or fistula. “Isolated Colonic IBD” included all patients with disease limited to the colon (724 with UC, 53 with IBD-U, and 402 with Colonic CD). A sub-group of IBD patients employed in this study (1,011 patients, including 647 CD and 317 UC and 47 inflammatory bowel disease type unclassified (IBD-U)), were utilized in a previous IBD GWA analysis reporting on two novel IBD loci on chromosome 20q13 and 21q22<sup>1</sup>; however, only novel and non-overlapping loci are being described in this manuscript (**Supplementary Table 2**). The control group was recruited by CHOP clinicians, nursing and medical assistant staff within the CHOP Health Care Network, which includes primary care clinics and outpatient practices. The control subjects did not have IBD or evidence of chronic disease based on self-reported intake questionnaire or clinician-based assessment. The Research Ethics Board of the respective Hospitals and other participating centers approved the study, and written informed consent was obtained from all subjects (or their legal guardians). Details of the ascertainment and characterization of the IIBDGC cohort (RC-CD2) were provided in the original scan and replication publications<sup>2-6</sup>

### Supplementary Results

We observed several additional loci at suggestive levels of significance ( $P < 1 \times 10^{-6}$ ) in our discovery scans of CD, UC, and IBD. The 15q22 locus, located near the *SMAD3* gene, nominally associated with IBD in the CHOP-based replication cohort, but failed to attain genome-wide significance in the discovery + replication meta analysis (**Supplementary Table 3**).

15q22 attained suggestive level of significance ( $P < 1 \times 10^{-6}$ ) in our combined IBD scan. SNP rs16950687 ( $P = 6.67 \times 10^{-7}$ , OR = 1.20 [1.12-1.29]) on 15q22 lies in an LD block containing the genes *SMAD3*, a TGF $\beta$  activated transcriptional modulator, and IQCH, a protein thought to have a regulatory role in spermatogenesis. rs16950687 nominally replicates in replication cohort RC1 ( $P = 0.019$ , OR = 1.21 [1.03-1.41]) but fails to replicate in the IIBDGC cohort (RC2-CD). rs16950687 also shows nominal association with CD in the majority adult onset meta analysis dataset ( $P = 0.0287$ , OR = 1.10) (**Supplementary Table 4**)<sup>2</sup>.

Our CD analysis showed an additional novel locus in a gene rich region of 1q22 (**Supplementary Table 3**). rs3180018 shows suggestive association in DC-CD ( $P = 7.76 \times 10^{-7}$ , OR = 1.25 (1.14-1.36)) but does not replicate in RC1-CD or RC2-CD. It however shows significance in the previously published majority adult-onset CD meta analysis ( $P = 0.02$ ,  $Z = 2.32$ ) (**Supplementary Table 4**), and thus may merit further followup as a CD risk variant<sup>2</sup>.

Our UC analysis revealed three additional novel loci (18q12.2, 16q21, 10q25.3) demonstrating suggestive levels of significance ( $P < 1 \times 10^{-6}$ ), however none of these successfully replicated in their respective replication cohorts (**Supplementary Table 3**). Given the small size of the RC1-UC dataset (120 UC cases, 1696 controls), further efforts to replicate the UC loci with better powered cohorts are warranted.

Finally, we observed significance at early onset loci near 20q13 (near *TNFRSF6B*) and 21q22 (near *BWRDI* and *PSMG1*) previously identified in a genome wide scan of a subset of the cohort used in this study<sup>1</sup>. Our current dataset further supported these associations: rs2315008 on 20q13 showed robust association with CD in the DC-CD, RC1-CD, RC2-CD meta analysis ( $P = 3.50 \times 10^{-9}$ ,  $Z = -5.91$ ) and genome wide significance in the combined IBD meta analysis (DC-CD + RC1-IBD, RC2-CD) ( $P = 4.67 \times 10^{-11}$ ,  $Z = -6.58$ ). rs2315008 showed less significant association in the UC-only (DC-UC, RC1-UC) meta analysis ( $P = 1.05 \times 10^{-4}$ ,  $Z = -3.88$ ). rs2836878 on 21q22 showed the most striking association with UC ( $P = 2.65 \times 10^{-9}$ ,  $Z = -5.95$ ) and nominal significance in the combined IBD meta analysis ( $P = 1.21 \times 10^{-6}$ ,  $Z = -4.85$ ). These results suggest that 20q13 may be a more CD specific locus while 21q22 is a more UC specific marker.

## Additional Authors

### Western Regional Alliance for Pediatric IBD

Gary Silber<sup>1</sup>, Iwona Wrobel<sup>2</sup>, Antonio Quiros,<sup>3</sup>

<sup>1</sup>Department of Pediatrics, Phoenix Children's Hospital, 1919 East Thomas Road, Phoenix, AZ 85016, USA

<sup>2</sup>Department of Pediatrics, Alberta Children's Hospital, 2888 Shaganappi Trail NW, Calgary, Alberta Canada T3B 6A8

<sup>3</sup>Department of Pediatrics, California Pacific Medical Center, 2340 Clay Street, San Francisco, CA 94120 USA

### International IBD Genetics Consortium

Jeffrey C. Barrett<sup>1</sup>, Sarah Hansoul<sup>2</sup>, Dan L. Nicolae<sup>3</sup>, Judy H. Cho<sup>4</sup>, Richard H. Duerr<sup>5,6</sup>, John D. Rioux<sup>7,8</sup>, Steven R. Brant<sup>9,10</sup>, Mark S. Silverberg<sup>11</sup>, Kent D. Taylor<sup>12</sup>, M. Michael Barmada<sup>5</sup>, Alain Bitton<sup>13</sup>, Themistocles Dassopoulos<sup>9</sup>, Lisa Wu Datta<sup>9</sup>, Todd Green<sup>8</sup>, Anne M. Griffiths<sup>14</sup>, Emily O. Kistner<sup>15</sup>, Michael T. Murtha<sup>4</sup>, Miguel D. Regueiro<sup>6</sup>, Jerome I. Rotter<sup>12</sup>, L. Philip Schumm<sup>15</sup>, A. Hillary Steinhart<sup>11</sup>, Stephan R. Targan<sup>12</sup>, Ramnik J. Xavier<sup>16</sup>, the NIDDK IBD Genetics Consortium, Cécile Libioulle<sup>2</sup>, Cynthia Sandor<sup>2</sup>, Mark Lathrop<sup>17</sup>, Jacques Belaiche<sup>18</sup>, Olivier Dewit<sup>19</sup>, Ivo Gut<sup>17</sup>, Simon Heath<sup>17</sup>, Debby Laukens<sup>20</sup>, Myriam Mni<sup>2</sup>, Paul Rutgeerts<sup>21</sup>, André Van Gossum<sup>22</sup>, Diana Zelenika<sup>17</sup>, Denis Franchimont<sup>22</sup>, JP Hugot<sup>23</sup>, Martine de Vos<sup>20</sup>, Severine Vermeire<sup>21</sup>, Edouard Louis<sup>18</sup>, the Belgian-French IBD consortium, the Wellcome Trust Case Control Consortium, Lon R. Cardon<sup>1</sup>, Carl A. Anderson<sup>1</sup>, Hazel Drummond<sup>24</sup>, Elaine Nimmo<sup>24</sup>, Tariq Ahmad<sup>25</sup>, Natalie J Prescott<sup>26</sup>, Clive M. Onnie<sup>26</sup>, Sheila A. Fisher<sup>26</sup>, Jonathan Marchini<sup>27</sup>, Jilur Ghorri<sup>28</sup>, Suzannah Bumpstead<sup>28</sup>, Rhian Gwillam<sup>28</sup>, Mark Tremelling<sup>29</sup>, Panos Deloukas<sup>28</sup>, John Mansfield<sup>30</sup>, Derek Jewell<sup>31</sup>, Jack Satsangi<sup>24</sup>, Christopher G. Mathew<sup>26</sup>, Miles Parkes<sup>29</sup>, Michel Georges<sup>2</sup>, Mark J. Daly<sup>8,32</sup>

<sup>1</sup>Bioinformatics and Statistical Genetics, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK.

<sup>2</sup>Unit of Animal Genomics, GIGA-R and Faculty of Veterinary Medicine, University of Lie`ge, Belgium.

<sup>3</sup>University of Chicago, Department of Medicine, 5801 South Ellis, Chicago, Illinois 60637, USA.

<sup>4</sup>Yale University, Departments of Medicine and Genetics, Division of Gastroenterology, Inflammatory Bowel Disease (IBD) Center, 300 Cedar Street, New Haven, Connecticut 06519, USA.

<sup>5</sup>University of Pittsburgh, School of Medicine, Department of Medicine, Division of Gastroenterology, Hepatology and Nutrition, University of Pittsburgh Medical Center (UPMC) Presbyterian, 200 Lothrop Street, Pittsburgh, Pennsylvania 15213, USA.

<sup>6</sup>University of Pittsburgh, Graduate School of Public Health, Department of Human Genetics, 130 Desoto Street, Pittsburgh, Pennsylvania 15261, USA.

<sup>7</sup>Université de Montréal and the Montreal Heart Institute, Research Center, 5000 rue Belanger, Montreal, Quebec H1T 1C8, Canada.

<sup>8</sup>The Broad Institute of Massachusetts Institute of Technology and Harvard, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA.

<sup>9</sup>Johns Hopkins University, Department of Medicine, Harvey M. and Lyn P. Meyerhoff Inflammatory Bowel Disease Center, 1503 East Jefferson Street, Baltimore, Maryland 21231, USA.

<sup>10</sup>Johns Hopkins University, Bloomberg School of Public Health, Department of Epidemiology, 615 E. Wolfe Street, Baltimore, Maryland 21205, USA.

<sup>11</sup>Mount Sinai Hospital IBD Centre, University of Toronto, 441-600 University Avenue, Toronto, Ontario M5G 1X5, Canada.

<sup>12</sup>Medical Genetics Institute and Inflammatory Bowel Disease (IBD) Center, Cedars-Sinai Medical Center, 8700 W. Beverly Blvd., Los Angeles, California 90048, USA.

- <sup>13</sup>Department of Medicine, Royal Victoria Hospital, McGill University, Montreal, Quebec, H3A 1A1, Canada.
- <sup>14</sup>The Hospital for Sick Children, University of Toronto, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada. <sup>15</sup>University of Chicago, Department of Health Studies, 5841 S. Maryland Avenue, Chicago, Illinois 60637, USA.
- <sup>16</sup>Gastrointestinal Unit and Center for Computational and Integrative Biology, Massachusetts General Hospital, Harvard Medical School, 185 Cambridge Street, Boston, Massachusetts 02114, USA.
- <sup>17</sup>Centre National de Génotypage, Evry, France.
- <sup>18</sup>Unit of Hepatology and Gastroenterology, Department of Clinical Sciences, GIGA-R, Faculty of Medicine and CHU de Liège, University of Liège, Belgium.
- <sup>19</sup>Department of Gastroenterology, Clinique universitaire St Luc, UCL, Brussels, Belgium. <sup>20</sup>Department of Hepatology and Gastroenterology, Ghent University Hospital, Belgium. <sup>21</sup>Department of Gastroenterology, University Hospital Leuven, Belgium.
- <sup>22</sup>Department of Gastroenterology, Erasmus Hospital, Free University of Brussels, Belgium. <sup>23</sup>INSERM; Université Paris Diderot; Assistance Publique Hôpitaux de Paris; Hôpital Robert Debré, Paris, France.
- <sup>24</sup>Gastrointestinal Unit, Division of Medical Sciences, School of Molecular and Clinical Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK.
- <sup>25</sup>Peninsula Medical School, Barrack Road, Exeter, EX2 5DW, UK.
- <sup>26</sup>Department of Medical and Molecular Genetics, King's College London School of Medicine, 8th Floor Guy's Tower, Guy's Hospital, London, SE1 9RT, UK.
- <sup>27</sup>Department of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK.
- <sup>28</sup>The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.
- <sup>29</sup>IBD research group, Addenbrooke's Hospital, University of Cambridge, Cambridge CB2 2QQ, UK.
- <sup>30</sup>Department of Gastroenterology and Hepatology, University of Newcastle upon Tyne, Royal Victoria Infirmary, Newcastle upon Tyne NE1 4LP, UK.
- <sup>31</sup>Gastroenterology Unit, Radcliffe Infirmary, University of Oxford, Oxford, OX2 6HE, UK.
- <sup>32</sup>Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, 185 Cambridge Street, Boston, Massachusetts 02114, USA.

## REFERENCES

1. Kugathasan, S. et al. Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. *Nat Genet* **40**, 1211-5 (2008).
2. Barrett, J.C. et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* **40**, 955-62 (2008).
3. Dixon, A.L. et al. A genome-wide association study of global gene expression. *Nat Genet* **39**, 1202-7 (2007).

## TABLES

### Supplementary Table 1

Study recruitment, subsequent inclusion, and ultimate demographic and phenotypic characteristics of Caucasian (European ancestry) subjects with matched controls in the discovery cohort DC-IBD (n=2413)

(a) General demographic characteristics of cohorts employed in our study cohort DC-IBD

		IBD	CD	UC	IBD-U
Recruited		3370	2304	993	73
Caucasian + meeting QC criteria		2784	1887	835	62
Final Discovery Cohort		2413	1636	724	53
Demographic Characteristics	Male	1273 (52.7%)	927 (56.6%)	321 (44.3%)	25 (47.2%)
	Median Age at Diagnosis (IQR)	12yrs (9-14.2)	12yrs (10-14)	12yrs (8-15)	10.25yrs (7-13.5)
	1° Familial Hx (Valid %) <sup>1</sup>	289 (14%)	215 (15.5%)	63 (10.2%)	11 (21%)
	Known Jewish Heritage (Valid %) <sup>2</sup>	223 (9.6%)	161 (10.3%)	57 (8.1%)	5 (9.8%)

(b) Specific disease characteristics of DC-CD and DC-UC discovery cohorts.

CD Patient Characteristics			UC Patient Characteristics		
Disease Behaviour <sup>6</sup>	Fibrosenotic	187 (15.7%)	Disease Behaviour <sup>6</sup>	Extensive Disease	394 (70%)
	Internally Penetrating	190 (15.9%)		Left-Sided Disease	168 (30%)
Anatomic Location <sup>3</sup>	Isolated Small Bowel Disease (Valid %)	297 (20%)			
	Isolated Colonic Disease (Valid %)	402 (27.2%)			
	Small Bowel Colon Disease (Valid %)	769 (52%)			
	Any Perianal Disease <sup>5</sup> (Valid %)	312 (21.4%)			

(c) Ethnic origins of our discovery cohort DC-IBD.

	IBD Discovery Cohort (DC-IBD)
Italy	322
Scotland	374
Canada	528
United States	1189
<b>TOTAL</b>	<b>2413</b>

1. Family Hx details not available in 14% of cases
2. Jewish Heritage unknown in 4% of cases
3. 7 cases had disease isolated to the upper tract, one case had disease isolated to the perianal region. Complete disease location data unavailable in 10% of CD cases
4. Details of disease extent unavailable in 22% of UC cases
5. Details of perianal disease unavailable in 11% of CD cases
6. Details of disease behaviour at latest review unavailable in 27% of CD cases

Supplementary Table 2

Discovery cohort (DC-IBD) sizes and filtering

	<i>Kuthagasan et al</i> <sup>1</sup>			Consortium			All			Controls
	CD	UC	IBD	CD	UC	IBD	CD	UC	IBD	
QC Filtered + Caucasian	647	317	1011	1241	548	1677	1887	835	2722	7315
Final Cohort	606	308	903	966	470	1510	1636	724	2413	6158

### Supplementary Table 3

- a) Novel genome wide significant ( $P < 1 \times 10^{-7}$ ) and suggestive ( $P < 1 \times 10^{-6}$ ) CD loci identified in the discovery GWA scan of early-onset CD patients (DC-CD). Since rs3180018 was not typed in the IIBDGC study (RC2-CD), we assessed association in this dataset at a nearby SNP rs1052176 (fields marked with a “\*”).

Band	MB	Genes	SNP	All	CD Discovery (n = 1636 / 6158)				CHOP CD replication (RC1-CD) (n = 289 / 1696)				IIBDGC replication (RC2-CD) (n = 531 / 4109)				Replication combined		Total Combined	
					P	Aff	Unaff	OR	P	Aff	Unaff	OR	P	Aff	Unaff	OR	P	Z	P	Z
16p11.2	28.45-28.54	IL27, SULT1A1, SULT1A2, EIF3C	rs1968752	A/T	2.09E-08	0.39	0.34	1.25 [1.16-1.36]	0.81	0.36	0.35	1.02 [0.85-1.23]	0.036	0.35	0.33	1.09 [0.94-1.27]	0.059	1.89	6.67E-08	5.40
1q22	153.46-154.06	C1orf2, CLK2, GBA, HCN3, SCAMP3	rs3180018	A/T	7.76E-07	0.28	0.24	1.25 [1.14-1.36]	0.10	0.28	0.25	1.17 [0.97-1.40]	0.91*	0.26*	0.25*	1.03 [0.89-1.19]*	0.412	0.82	2.89E-05	4.18

- b) Novel genome wide significant ( $P < 1 \times 10^{-7}$ ) and suggestive ( $P < 1 \times 10^{-6}$ ) UC loci identified in the discovery GWA scan of early-onset UC patients (DC-UC).

Band	MB	Genes	SNP	All	UC Discovery (DC-UC) (n=724 / 6158)				CHOP UC Replication (RC1-UC) (n=120 / 1696)				Total Combined	
					P	Aff	Unaff	OR	P	Aff	Unaff	OR	P	Z
18q12.2	32.22-32.25	FHOD3, MOCOS	rs7228236	C/G	9.93E-08	0.16	0.22	0.67 [0.58-0.78]	0.84	0.23	0.22	1.03 [0.76-1.41]	3.35E-06	-4.65
2q37.3	241.21-241.42	CAPN10, GPR35, KIF1A, RNPEPL1	rs4676410	A/T	1.70E-07	0.24	0.18	1.41 [1.24-1.61]	0.06	0.25	0.20	1.34 [0.99-1.82]	3.64E-08	5.51
16q21	57.06-57.07	NDRG4	rs16960173	A/T	5.67E-07	0.34	0.28	1.34 [1.20-1.51]	0.58	0.28	0.27	1.09 [0.81-1.45]	2.54E-06	4.70
10q25.3	115.17-115.26	HABP2, NRAP	rs12360212	A/T	8.55E-07	0.30	0.24	1.35 [1.20-1.52]	0.5139	0.21	0.23	0.90 [0.65-1.24]	4.50E-05	4.08

- c) Novel genome wide significant ( $P < 1 \times 10^{-7}$ ) and suggestive ( $P < 1 \times 10^{-6}$ ) IBD loci identified in the combined discovery GWA scan of early onset IBD (DC-IBD).

Band	MB	Genes	SNP	All	IBD Discovery (n = 2413 / 6158)				CHOP IBD replication (RC1) (n = 482 / 1696)				IIBDGC replication (RC2-CD) (n = 531 / 4109)				Replication combined		Total Combined	
					P	Aff	Unaff	OR	P	Aff	Unaff	OR	P	Aff	Unaff	OR	P	Z	P	Z
8q24.21	128.25-128.28		rs2456449	C/G	1.26E-07	0.30	0.34	0.82 [0.77-0.89]	0.30	0.32	0.33	0.92 [0.79-1.07]	0.32	0.33	0.34	0.94 [0.82-1.08]	0.16	-1.41	1.02E-06	-4.89
16p11.2	28.74-28.81	IL27, SULT1A1, SULT1A2, EIF3C	rs8049439	C/G	2.38E-07	0.41	0.37	1.20 [1.12-1.28]	0.34	0.40	0.36	1.08 [0.93-1.24]	0.0014	0.42	0.39	1.14 [1.00-1.30]	0.0015	3.17	2.41E-09	5.97
15q22.33	65.25-65.26	SMAD3	rs18950687	C/G	5.24E-07	0.31	0.27	1.20 [1.12-1.29]	0.019	0.31	0.28	1.21 [1.03-1.41]	0.26	0.27	0.26	1.06 [0.92-1.23]	0.024	2.26	1.51E-07	5.25
22q12.2	28.75-28.86	HORMAD2, MTRR3, LIF	rs2412973	A/T	9.14E-07	0.50	0.46	1.18 [1.11-1.26]	0.0052	0.51	0.46	1.23 [1.06-1.42]	0.016	0.50	0.46	1.15 [1.01-1.31]	3.58E-04	3.57	1.54E-09	6.04

## Supplementary Table 4

- a) Novel genome wide significant ( $P < 1 \times 10^{-7}$ ) and suggestive ( $P < 1 \times 10^{-6}$ ) putative CD loci identified in the discovery GWA scan of early-onset CD patients (DC-CD) and their corresponding significance (or that of a surrogate) in the previously published majority adult-onset CD meta analysis.

Band	MB	Genes	SNP	CD Discovery (n = 1636 / 6158)				CD meta analysis (n = 3230 / 4829)		
				P	Aff	Unaff	OR	SNP	P	Z
16p11.2	28.45-28.81	<i>IL27, SUL1A1, SUL1A2, EIF3C</i>	rs1968752	2.09E-08	0.39	0.34	1.25 [1.16-1.36]	rs4788084	0.0035	2.92
1q22	153.46-154.06	<i>C1orf2, CLK2, GBA, HCN3, SCAMP3</i>	rs3180018	7.76E-07	0.28	0.24	1.25 [1.14-1.36]	rs1052176	0.020	2.33

- b) Novel genome wide significant ( $P < 1 \times 10^{-7}$ ) and suggestive ( $P < 1 \times 10^{-6}$ ) putative IBD loci identified in the discovery GWA scan of early-onset CD patients (DC-CD) and their corresponding significance (or that of a surrogate) in the previously published majority adult-onset CD meta analysis.

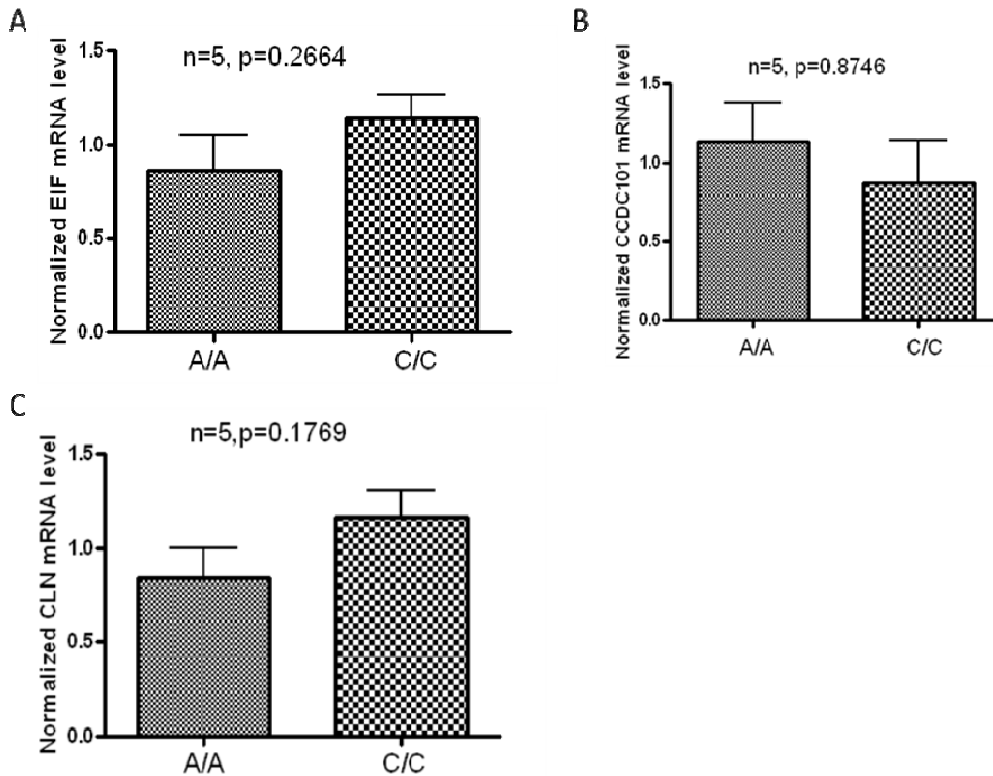
Band	MB	Genes	SNP	IBD Discovery (n = 2413 / 6158)				CD meta analysis (n = 3230 / 4829)		
				P	Aff	Unaff	OR	SNP	P	Z
8q24.21	128.25-128.28		rs2456449	1.26E-07	0.30	0.34	0.82 [0.77-0.89]	rs2456449	2.33E-01	1.19
16p11.2	28.45-28.81	<i>IL27, SUL1A1, SUL1A2, EIF3C</i>	rs8049439	2.38E-07	0.41	0.37	1.20 [1.12-1.28]	rs8049439	4.96E-03	2.81
15q22.33	65.25-65.26	<i>SMAD3</i>	rs16950687	5.24E-07	0.31	0.27	1.20 [1.12-1.29]	rs16950687	2.87E-02	2.19
22q12.2	28.75-28.86	<i>HORMAD2, MTMR3, LIF</i>	rs2412973	9.14E-07	0.50	0.46	1.18 [1.11-1.26]	rs2412973	9.53E-04	3.30

**Supplementary Table 5** 49 previously identified known and putative adult-onset IBD loci examined by our study evaluated in a meta analysis of our discovery (DC-CD, DC-UC, DC-IBD) and replication (RC1, RC1-CD, RC1-UC) cohorts. This analysis did not include the IIBDGC dataset (RC2-CD) which was a subset of the original majority-adult onset dataset which identified some of these loci. Filled circles in the first four columns of the table specify whether the given row represents a known CD locus (CDk), putative / nominal CD locus (CDp), known UC locus (UCk), and / or putative / nominal UC locus (UCp), respectively. We validate 23 of 32 known adult-onset CD loci, 8 of 17 known adult-onset UC loci, and overall 29 of 41 known and 5 of 10 putative adult-onset IBD loci in our early onset CD, UC, and IBD datasets. Loci demonstrating Bonferroni-corrected  $P < 0.05$  are denoted in bold (corrected for 49 hypotheses). Our data also implicate several previously described CD loci as having association with UC (bold italics). We also verify 2 nominally associating SNPs from the recent CD meta analysis (bold italics).

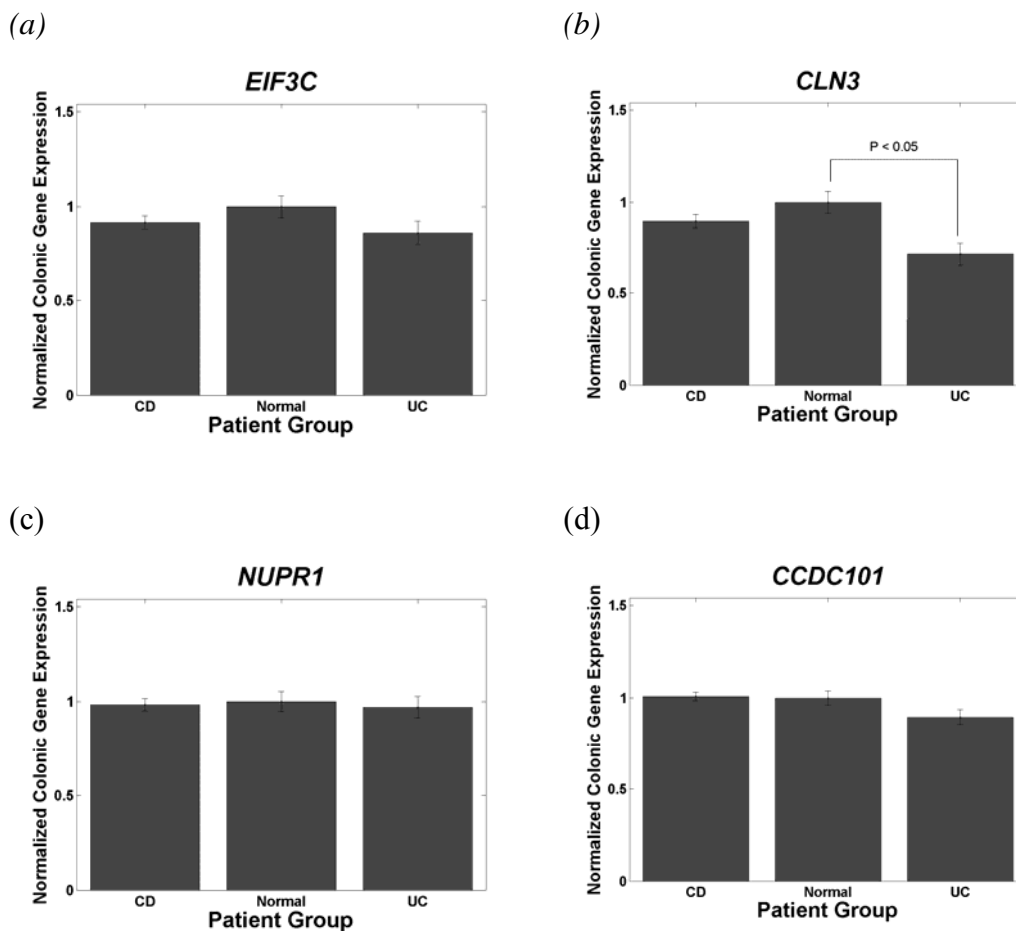
CDk	CDp	UCk	UCp	Band	MB	Genes	SNP	CD (DC-CD, RC1-CD)		UC (DC-UC, RC1-UC)		IBD (DC-IBD, RC1)	
								(n = 1925 / 7854)		(n = 844 / 7854)		(n = 2895 / 7854)	
								P	Z	P	Z	P	Z
•				1p13.2	114.18	<i>PTPN22</i>	rs2476801	<b>2.14E-06</b>	<b>-4.74</b>	4.92E-01	0.69	<b>1.86E-04</b>	<b>-3.74</b>
•	•			1p31.3	67.48	<i>IL23R</i>	rs11465804	<b>1.67E-15</b>	<b>-7.96</b>	2.92E-03	-2.98	<b>1.03E-15</b>	<b>-8.02</b>
	•			1p36.13	20.04	<i>PLA2G2E, OTUD3</i>	rs6426833	4.17E-01	0.81	<b>2.19E-06</b>	<b>-4.73</b>	4.69E-02	-1.99
		•		1q21.2	148.75		rs13294	7.67E-01	0.30	2.48E-02	2.24	1.70E-01	1.34
•				1q23.3	159.12	<i>OR10J1</i>	rs2274910	1.27E-01	-1.53	3.58E-01	-0.92	1.23E-01	-1.54
•				1q24.3	171.13	<i>FMO4</i>	rs9286879	<b>4.19E-06</b>	<b>4.60</b>	9.46E-01	0.07	<b>4.65E-04</b>	<b>3.50</b>
•				1q32.1	199.25		rs12122721	3.47E-01	-0.94	1.91E-01	-1.31	8.03E-02	-1.75
	•			1q32.1	205.01	<i>IL10, IL19, IL20</i>	rs3024505	<b>4.84E-04</b>	<b>3.49</b>	<b>6.20E-04</b>	<b>3.42</b>	<b>6.84E-06</b>	<b>4.50</b>
•				2p16.1	61.04	<i>AHSA2, CCDC139, PEX13, USP34, PUS10</i>	rs13003464	4.30E-03	2.88	5.51E-02	1.92	1.20E-03	3.24
•				2p23.3	27.59	<i>GCKR</i>	rs780094	1.45E-01	1.46	2.78E-03	2.99	3.60E-03	2.91
•				2q12.1	102.44	<i>IL18R1, IL18RAP</i>	rs917997	<b>6.84E-05</b>	<b>3.98</b>	1.67E-01	1.38	<b>5.98E-05</b>	<b>4.01</b>
	•			2q35	218.77	<i>Multiple</i>	rs6752254	4.29E-01	-0.79	3.53E-03	-2.92	3.23E-02	-2.14
•				2q37.1	233.85	<i>DOKD</i>	rs2241880	<b>1.09E-18</b>	<b>-8.83</b>	7.58E-01	-0.31	<b>9.46E-14</b>	<b>-7.45</b>
	•			3p12.1	85.84	<i>CADM2</i>	rs7611991	9.90E-01	0.01	1.22E-02	-2.51	1.13E-01	-1.58
•	•			3p21.31	49.70	<i>MST1</i>	rs3197999	<b>6.57E-10</b>	<b>6.18</b>	<b>4.31E-04</b>	<b>3.52</b>	<b>2.41E-11</b>	<b>6.68</b>
•		•		5p13.1	40.43	<i>PTGER4</i>	rs4613763	<b>8.98E-05</b>	<b>3.92</b>	3.52E-01	0.93	1.02E-04	3.89
	•			5q13.3	76.18	<i>F2RL1, S100Z</i>	rs7724915	3.93E-01	-0.85	6.65E-02	1.84	8.52E-01	0.19
•	•			5q31.1	131.80	<i>Multiple</i>	rs2188962	3.55E-08	5.51	6.29E-01	0.48	<b>3.28E-06</b>	<b>4.65</b>
•				5q33.1	150.25	<i>ZNF300</i>	rs7714584	<b>3.37E-04</b>	<b>3.59</b>	7.26E-02	1.80	<b>1.66E-04</b>	<b>3.77</b>
•	•			5q33.3	158.75	<i>IL12B, RNF145, UBLCP1</i>	rs10045431	<b>2.63E-06</b>	<b>-4.70</b>	<b>4.14E-07</b>	<b>-5.06</b>	<b>4.38E-11</b>	<b>-6.59</b>
•				6p21.32	32.54	<i>BTNL2, SLC26A3, HLA-DRB1, HLA-DQA1</i>	rs2395185	1.91E-01	-1.31	<b>7.40E-19</b>	<b>-8.87</b>	<b>7.70E-09</b>	<b>-5.77</b>
	•			6p21.32	32.69	<i>HLA-DRA</i>	rs660895	1.05E-02	-2.56	<b>4.76E-15</b>	<b>-7.83</b>	<b>7.29E-10</b>	<b>-4.16</b>
•				6p22.3	20.84	<i>CDKAL1</i>	rs6908425	1.00E-02	-2.57	3.81E-02	-2.07	6.93E-03	-2.70
	•			6p25.1	5.10	<i>LYRM4</i>	rs12529198	3.72E-01	0.89	8.36E-01	-0.21	5.00E-01	0.68
	•			6p25.2	3.38	<i>C6orf85</i>	rs4959832	2.06E-01	-1.26	4.26E-01	-0.80	3.37E-01	-0.96
•				6q21	106.58		rs6938089	9.49E-02	1.67	9.96E-01	0.00	1.57E-01	1.42
	•			6q25.1	149.62		rs7758080	3.25E-01	0.99	6.27E-01	0.49	2.89E-01	1.06
•				6q27	167.36	<i>CCR6, FGFR1OP, GPR31, RINASE2</i>	rs2301436	2.61E-03	3.01	7.05E-01	0.38	1.77E-02	2.37
•				7p12.2	50.24	<i>ZPBP</i>	rs1456893	<b>9.04E-06</b>	<b>-4.44</b>	5.82E-01	-0.55	<b>4.33E-04</b>	<b>-3.52</b>
•				8q24.13	126.61		rs1551398	<b>6.55E-07</b>	<b>-4.97</b>	6.15E-01	-0.50	<b>9.21E-05</b>	<b>-3.91</b>
•	•			9p24.1	4.97	<i>INSL6, JAK2</i>	rs10758669	<b>6.44E-06</b>	<b>4.51</b>	6.56E-03	2.72	<b>5.11E-07</b>	<b>5.02</b>
•				9q32	116.60	<i>SLC46A2</i>	rs6478108	<b>6.73E-09</b>	<b>-5.80</b>	2.88E-03	-2.98	<b>5.06E-10</b>	<b>-6.22</b>
•	•			10p11.21	35.43	<i>CCNY, CREM, CUL2</i>	rs4934724	<b>1.73E-05</b>	<b>4.30</b>	1.45E-01	1.46	<b>4.34E-06</b>	<b>4.59</b>
•		•		10q21.2	64.07	<i>ZNF365</i>	rs10995250	<b>1.77E-07</b>	<b>5.22</b>	4.12E-02	2.04	<b>4.50E-07</b>	<b>5.05</b>
•	•			10q24.2	101.28	<i>NKX2-3</i>	rs11190140	<b>2.40E-09</b>	<b>-5.97</b>	<b>1.78E-04</b>	<b>-3.75</b>	<b>2.74E-11</b>	<b>-6.66</b>
•				11q13.5	75.95	<i>C11orf30</i>	rs7130588	<b>5.45E-04</b>	<b>3.46</b>	1.14E-01	1.58	<b>2.88E-04</b>	<b>3.63</b>
•				12q12	38.67	<i>LRRK2, SLC2A13</i>	rs11174631	<b>1.39E-05</b>	<b>4.35</b>	6.25E-01	-0.49	1.05E-03	3.28
	•			12q15	68.79	<i>IL26, IL22, IFNG</i>	rs1558744	6.71E-03	2.71	<b>9.69E-04</b>	<b>3.30</b>	<b>2.10E-04</b>	<b>3.71</b>
•				13q14.11	43.36	<i>C13orf31, CCDC122, ENOX1</i>	rs3764147	<b>2.11E-07</b>	<b>5.19</b>	3.32E-01	0.97	<b>1.83E-05</b>	<b>4.29</b>
	•			15q13.1	26.20	<i>HERC2, OCA2</i>	rs1667394	5.96E-01	-0.53	2.55E-01	1.14	7.57E-01	0.31
•				16q12.1	49.30	<i>CYLD, NKD1, NOD2, SLIC1</i>	rs2066843	<b>3.96E-26</b>	<b>19.57</b>	5.55E-01	0.59	<b>3.55E-18</b>	<b>8.69</b>
	•			17q12	29.61	<i>CCL11, CCL2, CCL7</i>	rs991804	<b>2.31E-04</b>	<b>-3.68</b>	2.78E-02	-2.20	<b>6.64E-05</b>	<b>-3.99</b>
•				17q12	35.29	<i>ORMDL3</i>	rs2872507	<b>3.65E-04</b>	<b>3.56</b>	<b>7.62E-04</b>	<b>3.37</b>	<b>3.13E-06</b>	<b>4.66</b>
•	•			17q21.2	37.77	<i>STAT3</i>	rs744166	9.33E-03	-2.60	1.34E-01	-1.50	4.19E-03	-2.86
•		•		18p11.21	12.80	<i>PTPN2</i>	rs1893217	4.32E-03	2.85	2.48E-01	1.15	3.34E-03	2.94
	•			18q11.2	17.93		rs8098673	8.15E-02	1.74	1.58E-01	1.41	4.33E-02	2.02
•				19p13.3	1.08	<i>SBNO2</i>	rs2024092	3.04E-02	2.17	1.25E-03	3.23	<b>7.04E-04</b>	<b>3.39</b>
•				21q21.1	15.74		rs1736148	<b>7.59E-06</b>	<b>-4.48</b>	5.76E-01	-0.56	<b>1.01E-05</b>	<b>-4.41</b>
•				21q22.3	44.44	<i>ICOSLG1</i>	rs762421	<b>3.32E-07</b>	<b>5.10</b>	<b>2.54E-05</b>	<b>4.21</b>	<b>3.19E-10</b>	<b>6.29</b>

## FIGURES

**Supplementary Figure 1. LCL eQTL analysis of rs1968752.** Comparison of the A/A genotype and C/C genotype for rs1968752 in lymphoblastoid cell lines showed no allele-specific gene expression changes for (A) *EIF3C*, (B) *CCDC101*, (C) *CLN3*, three genes located near *IL-27* in the 16p11 locus, associated with Crohn's disease in our study. We did not assess allele specific gene expression of the two other genes in the LD block (*NUPR1* and *SULT1A1*), however these did not show allele-specific changes in gene expression in a publicly available database<sup>3</sup>. Error bars represent s.e.m.



**Supplementary Figure 2. Colonic expression of 16p11 genes in CD, UC, and normal patients .** The LD block on 16p11 harboring rs1968752 and rs8049439 associates with early onset CD and IBD in our analyses. We examined the expression of genes *IL27*, *CCDC101*, *CLN3*, *EIF3C*, *NUPR1*, *SULT1A1*, and *SULT1A2* between 11 normal, 30 early-onset CD, 10 early-onset UC samples, using one-way analysis of variance with Tukey-Kramer multiple-comparison correction and significance thresholds of  $P < 0.05$ ,  $P < 0.001$ , and  $P < 0.0001$ . Data for *IL27*, *SULT1A1*, and *SULT1A2* are shown in **Figure 2** and **Figure 3** of the main body of the manuscript. We did not observe significant colonic gene differences in CD vs normal for (a) *EIF3C*, (b) *CLN3*, (c) *NUPR1*, and (d) *CCDC101*, however *CLN3* showed a significant effect in UC vs normal ( $P < 0.05$ ). Error bars represent s.e.m.



**Supplementary Figure 3. Colonic expression of 22q12 and 2q37 candidate genes in CD, UC, and normal patients.** We examined colonic gene expression differences for (a) *HORMAD2* (22q12), (b) *LIF* (22q12) (c) *GPR35* (2q37) (d) *KIF1A* (2q37) (e) *RNPEPL1* (2q37). Error bars represent s.e.m.

